RESEARCH REPORT SERIES
(*Statistics #2003-05*)

The Use of Statistical Methods in the U.S. Census: <u>Utah v. Evans</u>

Patrick J. Cantwell, Howard Hogan,
and Kathleen M. Styles

Statistical Research Division
U.S. Bureau of the Census
Washington D.C. 20233

Report Issued: September 4, 2003

The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

The Use of Statistical Methods in the U.S. Census: Utah v. Evans

by

Patrick J. Cantwell, Howard Hogan, and Kathleen M. Styles
U.S. Census Bureau
Washington, DC 20233-9100

patrick.j.cantwell@census.gov

ABSTRACT: As in all recent censuses, the U.S. Census Bureau used statistical methods in the 2000 Census to account for missing or contradictory information concerning the number of people living in some identified housing units. These statistically adjusted counts were used for Congressional apportionment and redistricting. In 2002, the Supreme Court ruled that this limited use of statistics was both lawful and constitutional. This paper provides a context for that decision, by tracing the evolution of statistical methods in the U.S. census, and the accompanying litigation. It then presents the statistical and legal arguments and issues raised in Utah v. Evans.

KEY WORDS: Imputation; Congressional apportionment; Probability sampling; Census coverage.

1. INTRODUCTION

Census taking is both a civic ceremony and a scientific undertaking. Statistical science is used throughout the process to provide the nation with the most accurate measures of the population possible under the law and within the constraints of time, resources, and respondent burden. The use of statistics has grown steadily during the past seven censuses. While there may be limits to the role statistics can play in our civic ceremony, the available evidence indicates that the accuracy of census data has improved greatly. This increase in accuracy can be attributed both to increased Congressional support and funding, and to the power of statistical science. A nonstatistical census would be a very poor census indeed.

To apportion the U.S. House of Representatives and set new boundaries for seats in state legislatures and local governments, the total number of persons residing in each census block in the country is required; information on the characteristics of these people is needed for other legislative programs and initiatives. Thus, census taking must confront fundamental issues concerning what the nation means by accuracy: numeric or distributive (proportional); supporting small areas or larger areas. Numeric accuracy refers to how close the overall count of a particular geographic area or demographic group is to the "true" number of people who reside in that area or belong to that group. Distributive accuracy refers to how close the relative proportion or share of a geographic area or demographic group is to its true share *relative to other areas or groups*. A perfect census--one in which every resident is counted once and only once and is correctly located--would be both numerically and distributively accurate.

The decennial census has never been a simple physical headcount of inhabitants, but rather has routinely employed techniques designed to obtain and use reliable information from a variety of sources. From the very first census, enumerators have relied on information provided by a member of the household, rather than on the enumerator's direct observation. All decennial censuses have relied on the observation and response from heads of households and other knowledgeable sources. When a member of the household could not be found, census takers have historically relied on information from proxy respondents, such as neighbors, landlords, or postal workers. Enumeration by proxy was first formally authorized for the 1880 Census.

More explicit statistical methods--procedures based on formal theory derived through the laws of probability and the science of statistics--also have a long history. Indeed, the use of statistical methods in taking the United States census evolved with the development of the science of statistics. Several times, when faced with the decision to accept an inaccurate census count or statistically adjust the result, the Census Bureau has tried to increase accuracy, as it was understood at the time, through the use of statistical methods.

As the census is the basis for dividing political power among the 50 states through the apportionment of the U.S. House of Representatives, the use of statistical procedures such as probability sampling to improve the accuracy of census data has been challenged in the courts in recent years. After the 1980 Census, the State of Indiana unsuccessfully sued to overturn the use of count imputation that moved a seat in the U.S. House of Representatives to Florida (Orr v. Baldrige 1985). (Count imputation is a statistical method used to account for missing or contradictory information concerning the number of residents in a housing unit.) Since 1980 many cases have been brought addressing the use of statistical adjustment to account for the differential coverage among various racial and demographic groups in the census. Leading up to the 2000 Census, the Census Bureau proposed a plan to sample nonresponding households to reduce cost, and to measure and potentially adjust for any net undercount. This plan was rejected by the United States Supreme Court in 1999, Dept. of Commerce v. U.S. House of Representatives (1999), as a means of obtaining the apportionment counts.

With sampling for nonresponse and adjustment prohibited for apportionment, the main legal and policy issue in the 2000 Census was expected to be whether adjusted census numbers should be used for nonapportionment purposes, such as state and local redistricting and the distribution of federal funds. Instead, while the adjustment issue was litigated extensively, the most significant lawsuit based on the 2000 Census was brought by the State of Utah challenging the use of count imputation. Utah claimed that the procedure was prohibited by both the Census Act and the Constitution. (The Census Act, 13 U.S.C. Section 1 *et seq*., establishes the Census Bureau and sets forth the Bureau's authority to conduct its business.) The Supreme Court resolved this issue in 2002, ruling in a split decision that count imputation was a permitted statistical method, different from sampling, Utah v. Evans (2002).

The Supreme Court's approval of at least limited use of statistical methods in Utah v. Evans was an important acceptance of the use of statistics to improve census accuracy. This acceptance clearly reflects the evolution of legal theory to accommodate modern scientific methods. One challenge in the years to come will be determining which statistical methods will pass the Supreme Court's acceptance test, and which should be considered barred under the Census Act or the Constitution.

2. THE USE OF STATISTICAL METHODS IN THE CENSUS BEFORE 2000

2.1 Early Censuses

Even with the first census in 1790, the nation faced the issue of how to handle missing census information. In that year, census returns from South Carolina were incomplete and Congress began to consider apportionment without them. The use of estimation was more explicit in 1850, when census returns for several California counties were lost at sea. The Superintendent of the Census, Joseph Kennedy, was faced with the issue of missing and incomplete data. According to the official publication, *Abstract of the Seventh Census* (1852),

> the population of California is set down at 165,000, as an approximation to the real population, which may be essentially varied by complete returns. Should the returns vary from ou[r] estimate so far as to reduce the population of California 30,000, South Carolina will be entitled to a member additional, as being next above on the list of fractions. (p. 153)
>
> Assuming the population of California to be 165,000, (which we do partly by estimate,) ... (p. 129)

Congress used the estimated California number for re-apportionment, choosing an estimate to account for missing information in an attempt to increase census accuracy (Anderson 1988, p. 46). Interestingly, the Bicentennial Edition *Historical Statistics of the United States*, shows the 1850 California population to be 93,000 (U.S. Census Bureau 1975, Series A 195-209, p. 25), ignoring Superintendent Kennedy's correction.

A different situation occurred in the 1920s. The 1920 Census was the first U.S. census to enumerate more people living in urban areas than in rural areas. Coincidentally or not, it is also the only census that might be called a failed census, insofar as it is the only one not used to re-apportion Congress (Anderson 1988). As a consequence, the apportionment devised after the 1910 Census was used for over 20 years, until after the completion of the 1930 Census. Although political considerations and a fight over the method of apportionment contributed toward the impasse, part of the Congressional discussion involved census accuracy. A major concern at the time was that Census Day had been set as January 1 rather than April 15 as in earlier censuses (or April 1, for those that followed). Congressmen from rural states alleged, with some justification, that census takers had difficulty canvassing rural areas during the winter months. In any case, discussions of census accuracy played a large role in the deliberations.

2.2 Statistical Science Begins to Have a Formal Role

The period from 1940 through 1990 was one of marked change in the census, with statistical methods increasingly used to improve quality while reducing cost and respondent burden. The 1940 Census initiated the use of sampling to simplify the collection of data on personal characteristics, while maintaining accuracy. In 1940 and 1950, additional detailed questions were asked of a sample of individuals within a household. Investigation of the results showed a subtle but measurable selection bias of individuals with respect to their relationship to the head of the household. By 1960, this would lead to the sampling of households rather than people. Then, a separate "long form" was used to collect more detailed information from the households selected in the sample.

With the exception of the introduction of sampling for person characteristics, the essential nature of census field operations remained similar from 1830 through 1950. Interviewers went door to door, listing housing units and interviewing the residents. The 1960 Census marked a genuine turning point with the switch to self-enumeration. For the first time written questionnaires were mailed to the respondents, who were asked to fill them out and hold them for pick-up. This was also the first census in which the apportionment count was tabulated by computers rather than by hand.

The 1970 Census was the first census in which the Census Bureau used the mail-out/mail-back procedure; the Bureau mailed out questionnaires, which were then completed and mailed back by respondents. From 1970 forward, this procedure has accounted for the largest share of the enumeration. To enable the mail-out, the Census Bureau developed a mailing list prior to Census Day; previously, enumerators had created a list of housing units as part of the interviewing process. With this address list, the Bureau could check and verify addresses prior to Census Day to achieve greater accuracy.

In the 1970 Census, two unusual situations arose that prompted the Census Bureau to apply imputation techniques in conjunction with sampling to add persons to the apportionment count.

Following the 1970 Census, the Bureau conducted two studies to measure the number of people missed in different situations. The National Vacancy Check was designed to measure and decrease the error introduced by misclassified housing units--occupied housing units that are incorrectly classified as vacant, or vice versa. The goal was to estimate the difference between the number of people missed because the census interviewer had incorrectly classified their housing unit as vacant, and those erroneously included although their housing unit was in fact vacant on Census Day. A second operation, the Post-Enumeration Post-Office Check (PEPOC), looked for missed housing units in selected areas of 16 Southern states. In each study, data were collected from a probability sample. When both studies found substantial numbers of people missed by the initial enumeration, statistical models were used to adjust the census before the numbers were reported for use in reapportionment.

The National Vacancy Check added 1,074,523 person records to the 1970 Census, constituting 0.53% of the total population count (U.S. Census Bureau 1971a). Because the National Vacancy Check was used across the nation, the differential impact among states was lessened. The additions by state ranged from 0.39% (Minnesota and Ohio) to 1.07% (Maine). The PEPOC added 589,517 person records to the 1970 Census, constituting 0.29% of the population. PEPOC, however, was conducted only in the 16 Southern states. The additions by state ranged from 0.32% of the population in Maryland, to 1.65% in North Carolina (U.S. Census Bureau 1971b). Between the two programs, North Carolina's population was adjusted upward by 123,575, or 2.4% of the total state population.

In each case, the Census Bureau *first* used sampling techniques to estimate the number of households missed, and *then* applied imputation techniques to approximate the population count and the characteristics in those households. Neither the National Vacancy Check nor the PEPOC was repeated on a sample basis in subsequent censuses for determining the apportionment counts, though several varieties of imputation (without the use of sampling) have consistently been used.

2.3 Administrative Records

In addition to statistical methods, the Census Bureau has often gone well beyond a simple head-count in an attempt to increase accuracy. City, state, and Federal records were used in the 1980 Census to enumerate people in a program known as the Non-Household Sources Program (U.S. Census Bureau 1984). This program was aimed at preventing the omission of urban minorities. It used a variety of administrative sources, including lists from state departments of motor vehicles and city public assistance files. Files from the U.S. Immigration and Naturalization Service were also used. These records were first clerically matched to census reports for the housing unit to identify people who may have been missed. Local census enumerators conducted telephone or personal interviews at these addresses to ensure that the person lived there on census day. If the person was reported to have lived there, he or she was added to the census.

In the 1990 Census, administrative records played an important role in two programs again designed to reduce the omission of minorities. Both programs involved people on parole or probation. In the first program, corrections officers were asked to give a special census form to the parolee when he or she came into the parole office for a routine visit. The parolee was asked to fill out the form, which the corrections officer returned to the Census Bureau. If the parolee/probationer was not counted at the address reported as his residence on April 1, he or she was added to the census. Unlike 1980, there was no follow-up at the reported residences (U.S. Census Bureau 1992a).

A second program was known as the Parolee/Probationer Coverage Improvement Program Follow-Up. In this program, administrative lists of the states' departments of corrections were accepted. These lists had to contain names, addresses, and at least two demographic characteristics for each inmate. Addresses on the lists were then certified by the departments of corrections as the parolees' or probationers' residences on Census Day, April 1, 1990. Direct contact with parolees and probationers was not required, nor was there follow-up at the reported addresses. Again, if the parolee or probationer was not listed at the Census-Day address reported, he or she was added to the census.

Administrative records were used most directly in the 1990 and 2000 Censuses to enumerate the military and federal employees living overseas, as well as their dependents. The enumeration was accomplished directly from administrative records with the numbers added to the resident enumeration. The overseas population was excluded from all tabulations except Congressional apportionment. This practice was upheld by the Supreme Court in 1992, Franklin v. Massachusetts (1992).

2.4 Considering Statistical Adjustment to Improve Coverage

After World War II, demographers and statisticians began to develop methods to estimate census coverage of the population. These methods verified what had long been suspected: that the decennial census misses some individuals, thereby undercounting the population. This undercount was believed to be differential, with certain groups being undercounted at a higher rate than others. Price (1947) compared the 1940 Census counts for men aged 21 to 35 to totals from the First Selective Service Registration on October 16, 1940. Based on reasonable assumptions, he estimated that the census covered about 97% of this age/sex group. But a similar examination restricted to Black men aged 21 to 35 yielded an estimate of only about 87% coverage by the census. By studying cohorts from the censuses of 1930, 1940, and 1950 by age, sex, and race, Coale (1955) drew similar conclusions concerning the differential undercoverage of the U.S. population by race. An interesting and more complete exposition of the issues surrounding undercoverage and adjustment to the census can be found in Anderson and Fienberg (1999). As the implications of these studies became known, they were followed by calls to apply methods to correct for the measured errors in the population count.

In general, there are two types of coverage errors: omissions and erroneous inclusions. The numerical difference between the two is defined as the net coverage error. People are sometimes counted, but in the wrong place. For example, if an individual is omitted where he or she actually lives, but erroneously included in a different county in the same state, that individual would be considered an omission for one county count, an erroneous enumeration in the count for the other county, but a correct enumeration for the state total. There would be coverage error at the county level, but not at the state level.

The net undercount is the difference between the true population and the tabulated census count. Undercount can also arise from misclassification error. For example, if a person who was actually 66 years of age was enumerated as 64, the person would contribute to an undercount of the number of people aged 65 and above, and an overcount of people under 65. Again, if subtotals are summarized over age, the difference disappears.

Demographic analysis is one method used to estimate the net undercount. By analyzing birth records, death records, information on immigration and emigration, and selected data from the current and past censuses, it is possible to construct an estimate of the true population. Demographic analysis has been done separately by age and sex, as well as by race (Black/Non-Black). Although demographic analysis estimates were first made in conjunction with the 1950 Census, reliable demographic estimates first became widely available following the 1970 Census. These early estimates clearly documented a *differential* net undercount: the net undercount for Blacks was higher than that for non-Blacks.

Starting in 1950, the Census Bureau has often tried to measure a census's coverage by conducting and evaluating a sample survey, typically known as a post-enumeration survey. Over the decades the method has been refined. Following the 1980 Census, the results of the survey were not available until many months after the legal deadline for completing the census and were subject to a number of limitations. Nonetheless, the possibility of applying the survey results to adjust the 1980 Census became subject to litigation.

The Census Bureau was sued in over 50 cases following the 1980 Census. The plaintiffs in some of these cases sought to use the results of demographic analysis to correct the census results. Essentially, the plaintiffs requested that the census totals for all areas be adjusted upward for Blacks and non-Blacks proportionally to the measured net undercount for that group, using a technique known as synthetic estimation. The Census Bureau responded that estimates from demographic analysis were too inaccurate for this use, and that a synthetic estimate based on only two groups was too crude to be useful.

While the Census Bureau's decision not to adjust the 1980 Census was upheld by the courts, the organization began an extensive review of its coverage measurement program to determine whether it was possible to adjust the 1990 Census. The goal was to produce a more accurate measure of the net undercount, and to produce it sooner. In 1987 the Department of Commerce directed the Census Bureau not to proceed with plans to produce adjusted census counts, prompting the filing of a lawsuit against the Department. An agreement prior to the 1990 Census required the Census Bureau to undertake a post-enumeration survey after the 1990 Census and announce by July 15, 1991 whether the results of that survey would be used to adjust the previously announced results. On July 15, 1991, the Secretary of Commerce announced that the 1990 Census would not be statistically adjusted. This decision was litigated for several years, culminating in a 1996 Supreme Court decision upholding the Secretary's decision against adjustment (Wisconsin v. City of New York 1996).

3. THE 2000 CENSUS: THE PLAN AND THE REALITY

3.1 The 2000 Census: Early Planning

After the 1990 Census, the Census Bureau, under the direction of Congress and with input from the National Academy of Sciences, various advisory groups, and stakeholders, began to review every aspect of the census to determine how to proceed with the 2000 Census. Many alternatives were considered, including a sample census, a census based on administrative records, and other options. The Census Bureau's initial plan involved the use of sampling to control costs and to improve accuracy by supplementing--but not replacing--initial comprehensive attempts to locate and enumerate every individual and housing unit. The adopted plan applied sampling in three programs: the follow-up of nonrespondents, a check on units designated by the U.S. Postal Service to be vacant, and a post-enumeration survey designed to measure and adjust for under- or overcoverage. The post-enumeration survey was to be known as the Integrated Coverage Measurement program or simply the Quality Check. This survey was designed to deliver statistically adjusted population numbers in time for use in reapportionment. In communicating its plans, the Census Bureau stated, "In 1994 the Department of Justice (DoJ) reviewed the Census Bureau's preliminary plans to use sampling in Census 2000 and issued a written opinion confirming that the plan was neither illegal nor unconstitutional" (U.S. Census Bureau 1997).

In planning the 2000 Census, the Census Bureau faced a great challenge. There had been a steady rise in the cost of conducting the census from decade to decade, even when that cost was computed in constant dollars and allowance was made for the increasing population. In constant 1990 dollars, the cost per household of the four censuses from 1960 to 1990 was estimated at $9, $11, $20, and $25, respectively (National Research Council 1995, p. 48). Despite this increase in cost, a differential net undercount across races and other demographic groups persisted.

The Census Bureau proposed the use of sampling to reduce the number of household visits to get an interview. Specifically, the Census Bureau would mail questionnaires to all known residential addresses. Through advertising and outreach, it would encourage all households to mail back their questionnaire. However, rather than visiting all addresses that did not return a questionnaire by mail, the Census Bureau would follow up only a sample in each census tract--an area averaging about 1,700 housing units--using statistical inference to account for the others. In a similar manner, instead of verifying that every housing unit reported to be vacant by the U.S. Postal Service was indeed vacant, only a sample of such units would be checked.

The Census Bureau saw two advantages to sampling for nonresponse. The first was financial: roughly 42 million nonresponding housing units would not require personal visits, appreciably reducing costs for hiring and training field workers. The second anticipated advantage was an improvement in the quality of the data collected. Sampling introduces a controlled, measurable random variation. The smaller workload could provide better field control, thereby reducing nonsampling errors--errors caused by all sources other than sampling, such as errors due to misunderstanding a census question, placing a person at the wrong address, or faulty transfer of the data from the form to electronic files. Further, census field work could be completed earlier, closer to the census reference date, April 1. The hope was that, for small areas such as tracts, the added sampling error would be balanced by reduced nonsampling errors. For larger areas, the plan was seen as providing better data as the relative sampling error decreases as the size of the area increases, while the nonsampling errors do not increase (Hogan and Waite, 1998).

An interesting but different facet of the original design for the 2000 Census was the Census Bureau's plan for counting people without a usual residence, a group that includes many of those considered to be homeless, through an operation called the Service Based Enumeration (U.S. Census Bureau 2000a). Although the number of people involved in this program was not large, the approach is an interesting use of statistics to measure a population that is difficult to enumerate using traditional methods. Special procedures were planned at homeless shelters, soup kitchens, mobile food vans, and certain outdoor locations. Because the procedures could only account for people at these facilities on the day of enumeration, the Bureau planned to apply a technique called multiplicity estimation to account for people who sometimes use these facilities, but did not use them on the enumeration day.

People who were enumerated were asked how many times in the past week they had used the facility. For example, a person who used the service only two nights out of the seven nights in the week leading up to the enumeration would be given a weighting factor of 7/2. The idea behind multiplicity estimation is that, under specified assumptions, such a person had only two chances in 7 to be found at the facility on the date selected for enumeration. Conversely, a person who used a service all seven nights that week would be given a weighting factor of 7/7 or 1, as he or she would have been enumerated no matter what day had been selected for enumeration. By applying the appropriate weight, those who are found at the service facilities on the evening of enumeration are weighted upwards to account for people who sometimes use services, but did not on that day.

Serious problems afflicted the multiplicity estimator in the 2000 Census. The usage question suffered a high level of nonresponse, and worse, apparently a very high level of response error. The Census Bureau feared that the increased coverage obtained by applying multiplicity estimation might be outweighed by the bias induced by inaccurate answers as well as the added variance. Thus, a decision was made not to apply multiplicity estimation in the Service Based Enumeration (U.S. Census Bureau 2001a). The unadjusted census count for this population then included only those people actually enumerated (after removing duplicates) at these service locations on the one day.

The Integrated Coverage Measurement, on the other hand, was the more controversial of the sampling procedures planned for the 2000 Census. It involved interviewing a massive sample of housing units, approximately 750,000, selected to include blocks from every area in the country representing all races and ethnic groups. The large sample size would allow the production of independent population estimates for each state, so that data from one state would not be used to determine the population total in another state. Dual-system estimation, a traditional method originating from the capture-recapture technique often applied to estimate the size of wild-life populations (Sekar and Deming 1949; Hogan and Wolter 1988), would be used to produce adjusted counts of people within larger demographic groups and geographic areas. Meanwhile, synthetic estimation would provide adjusted counts at lower geographic levels, such as census blocks.

In January 1999, the Supreme Court ruled that the Census Act precluded the use of sampling to arrive at the apportionment count, <u>Dept. of Commerce v. U.S. House of Representatives</u> (1999). As a result, the Census Bureau was forced to eliminate sampling programs from the 2000 Census. As these methods had formed the backbone of the original census plan, the Bureau had to resort to an entirely new design, based on its contingency plan.

3.2 The Accuracy and Coverage Evaluation

The Supreme Court's 1999 decision barred the use of the Integrated Coverage Measurement for apportionment, requiring that the post-enumeration survey be quickly redesigned. The revised survey, renamed the Accuracy and Coverage Evaluation (A.C.E.), was sharply cut in size to about 300,000 housing units. Its intent was to produce adjusted population numbers in time for their possible use for redistricting and the distribution of federal funds. Redistricting refers to the division of each state or local government area into voting districts and is distinguishable from apportionment, by which the 435 Representatives of the United States Congress are allotted to the 50 states based on the states' total population counts.

Based on results of the A.C.E. and through the use of dual-system estimation, the Census Bureau estimated the number of people missed or enumerated erroneously in the 2000 Census. By applying synthetic estimation, the Bureau estimated counts for census blocks and higher geographic levels. In addition to the steps necessary to produce adjusted population numbers, the Bureau instituted a program to assess the relative accuracy of the adjusted and unadjusted numbers--analyzing, synthesizing, and comparing the effects of errors on the two sets of counts.

When data from the A.C.E. had been processed and tabulated, the Department of Commerce had only a relatively brief period of time before making a decision to release unadjusted or adjusted numbers to the states for redistricting. To help make an informed decision, Census Bureau staff computed and studied the adjusted estimates, as well as the various component activities that contributed to these numbers, and summarized their findings in a series of reports directed to a committee of senior executives at the Bureau. The reports are available on-line at http://www.census.gov/dmd/www/ReportRec.htm.

On March 1, 2001, the Census Bureau's Deputy Director, in consultation with a committee of senior executives, recommended to the Department of Commerce that "the unadjusted census data be released as the Census Bureau's official redistricting data." The chief reason stated was that there was no conclusive information implying that either set--the unadjusted or the adjusted data--was more accurate for redistricting. Although the committee found that "both the census and the A.C.E. were efficient and effective operations that produced high quality data," they were concerned about inexplicable differences between estimates derived from demographic analysis and the A.C.E., as well as potential errors arising in the operations leading into the estimation (U.S. Census Bureau 2001b, Recommendation).

Subsequent months provided the Census Bureau additional time to study and compare the unadjusted and adjusted sets of data. After reviewing various evaluations and analyses, the executive committee concluded that "most Census 2000 and A.C.E. operations were of high quality ... [and] that improvements were achieved over both the 1990 census and the 1990 coverage measurement survey." However, new evidence suggested that the A.C.E. "overstated the net undercount by at least 3 million persons, and that the cause of this error was the A.C.E.'s failure to measure a significant number of census erroneous enumerations, many of which were duplicates" (2001c, p. i). For more information, a series of summary reports can be found on-line at http://www.census.gov/dmd/www/EscapRep2.html.

3.3 The Use of Statistical and Other Methods in the 2000 Census

Statistical methods underlay every aspect of the 2000 Census, from building the list of addresses and removing duplicate responses, to quality assurance and measuring coverage. The statistical processes applied in the 2000 Census included the primary selection algorithm, disclosure avoidance, the duplicate housing-unit operation, characteristic imputation, count imputation (discussed in detail in Section 4), and the Accuracy and Coverage Evaluation. In addition, other activities that did not involve the use of statistical methods were planned and conducted in an attempt to accurately enumerate the U.S. population. Some of these are described below.

In 2000, the Census Bureau went to great lengths to obtain a census response from every household, instituting new procedures and programs to improve the mail response rate and to provide other convenient ways to respond. The Bureau designed and implemented marketing and partnership programs to increase awareness of the decennial census and the need for public cooperation. The marketing program was designed around the first-ever paid advertising program, which included a national media campaign to increase mail response, advertising directed at historically undercounted populations, and special messages aimed at hard-to-enumerate people. In the partnership program, the Census Bureau worked with state and local partners nationwide to encourage all individuals to respond to the census.

Further, in the 2000 Census, for the first time people were allowed to respond to the census in a variety of ways. In addition to the most common method--filling out and returning the paper questionnaire--people could answer by telephone. If they wished, respondents could call Questionnaire Assistance centers to obtain help with their forms and to complete their response over the telephone. Through the Be Counted program, people could obtain a blank form available in various public locations, return it, and be included in the census. Some individuals responded to the census through the Internet.

These various programs and options--while increasing coverage--increased the likelihood of an additional problem: multiple responses from the same household. For example, one spouse might mail back the questionnaire, while the other called the Census Bureau. The solution was to remove duplicate responses, a much more difficult process than it may appear. To make sure individuals were not counted more than once, the Bureau developed and improved methods to match households and individuals.

Through a procedure called the primary selection algorithm, based on a set of pre-defined criteria, one of the several matched responses was designated as the "primary," the response to be included in the apportionment count. Individuals in the duplicated responses who were not identified on the primary might be added to the primary to enhance the coverage. All others were removed or "unduplicated." The greater variety of options for responding to the Census were not possible before 2000 because the statistical matching procedures were not adequate to permit the proper unduplication of responses, and the names of residents were generally not captured for computer searching and matching.

Procedures to review and insure quality have long been a part of the census. Formal quality control procedures were introduced in the 1960 Census (U.S. Census Bureau 1967). Since that time, the role of statistical quality control and assurance continued to grow and evolve, expanding to include listing, interviewing, keying, coding, procurement, and especially printing. The applications of quality control, based on sound statistical theory, including sampling and estimation, have been central to the improvement of the census. Major operations of the 2000 Census plan were subjected to quality assurance activities designed to detect and correct errors before they affected accuracy or data quality. These quality assurance programs used statistical methods to detect possible problems with census enumerators and to target problem cases for re-evaluation.

A number of additional operations were implemented to ensure as complete coverage as possible in the census enumeration. Computer edits were performed on mail-return questionnaires to identify returns that contained missing persons, missing personal information, or large housing units (more than six persons). Interviewers then conducted telephone interviews with those housing units during the Coverage Edit Follow-Up operation to obtain accurate data about the persons residing there. Another operation, Coverage Improvement Follow-Up, was conducted after the nonresponse follow-up. This operation required that an interviewer recheck housing units classified as vacant or nonexistent during the follow-up to ensure that no units were misclassified, as well as to check on new construction, lost or blank forms, and other coverage problems. Through this operation, 2.3 million occupied housing units were enumerated. Finally, after data collection efforts were completed, the data were processed through a number of computer operations to edit inconsistent or missing responses.

Another use of statistical methods was the Housing Unit Duplication Operation. This program was created and conducted during the census itself, after the Census Bureau discovered that its address list contained a relatively large number of duplicate addresses (U.S. Census Bureau 2000b). The Bureau concluded that the census address file might contain a significant number of duplicated person records, many of which were assigned to these duplicated addresses. The goal was to increase accuracy by eliminating the person duplication caused by these duplicate housing units. Applying statistical procedures, the Bureau designed a computer program to identify and remove from the housing-unit list duplicate addresses likely to contain duplicated person records. Using the results of the program and analyzing which addresses were likely to include people already listed, 1.4 million addresses were removed from the census, together with the corresponding 3.6 million person records (U.S. Census Bureau 2000c).

4. THE USE OF IMPUTATION IN THE CENSUS

The Census Bureau has long used both characteristic and count imputation to account for missing data. The phrase missing data refers to a wide variety of situations. In the 2000 Census, situations that resulted in missing data included incomplete or unavailable responses from housing units with previously confirmed addresses, conflicting data from the same housing unit, and failures in the data-capture process. The various types of missing data included characteristic data (information about an enumerated person, such as sex, race, age) and population count data (information about the number of occupants in an identified housing unit).

In the 2000 Census the Census Bureau processed data for over 120 million households, including over 147 million paper questionnaires and 1.5 billion pages of printed material. The workload during peak operations was about 3.3 million forms per day. Given the size of this undertaking, some amount of missing data was unavoidable.

The categories of missing data and the processes that lead to them have changed over the decades as the census has become increasingly computerized and centralized. In the 1940 Census, the Census Bureau introduced characteristic imputation simply to impute missing ages; the population count was not affected. For example, if a census response did not provide ages for the individuals living in the housing unit, but supplied all other information, age was imputed for those individuals. Characteristic imputation has been used in every census since 1940 and its use has not been challenged.

Count imputation, in contrast, assigns a population count to a housing unit. This technique is used when the Census Bureau is unable to secure any information regarding a given address, or when the Census Bureau has limited or contradictory information about the address but not the definitive number of occupants. As in prior censuses, the Census Bureau used count imputation in three instances in the 2000 Census:

- *Household Size Imputation.* When Census Bureau records indicated that a housing unit was occupied, but did not show the number of residents, a population count for the unit was imputed.

- *Occupancy Imputation.* When Census Bureau records indicated that a housing unit existed but not whether it was occupied or vacant, occupancy status (occupied or vacant) was imputed; then, if the unit was imputed to be occupied, a household size was imputed, as above.

- *Status Imputation* (referring to housing-unit status). When Census Bureau records had conflicting or insufficient information about whether an address represented a valid, non-duplicated housing unit, the status of the unit (occupied, vacant, or nonexistent) was imputed; then, if the unit was imputed to be occupied, household size was imputed, as above. Before applying status imputation to a unit, evidence of the physical existence or validity of the unit was *required*.

It should be noted that the housing-unit status "nonexistent" includes more that simply a vacant lot. Other possibilities include buildings used only for business purposes and structures that are not (or not yet) fit for habitation.

Since the 1960 Census, the Census Bureau has typically used hot-deck methods to impute counts. Hot-deck procedures use contemporaneous data, often from neighboring housing units, and preserve a realistic distribution of data, unlike cold-deck imputation, which uses information from a prior census or some other outside source.

4.1 Count Imputation in Prior Censuses

The 1960 Census was the first census to use computers to produce the apportionment count. It is then not surprising that it was also the first census to use count imputation to resolve occasional discrepancies between expected and actual numbers of computer records. Some of the missing household data resolved through count imputation in the 1960 Census appear to have been the result of mechanical difficulties, specifically problems with FOSDIC (Film Optical Sensing Device for Input to Computers), the device used to read filled-in circles on questionnaires. In general, a sequential hot-deck method was used: the missing data were taken from the previously processed census form, which usually represented a neighboring housing unit.

In 1960, the Census Bureau reported that "[p]ersons substituted" due to "noninterview" or "mechanical failure" were responsible for 0.5% of the total U.S. census count (U.S. Census Bureau 1965, p. 83). Because the Bureau's term "substitution" has at various times referred to imputation of a full set of characteristics (age, sex, etc.) for a known, enumerated individual, as well as to count imputation, it is unclear whether the numbers in the cited table refer just to the number of count imputations or include other forms. Moreover, the Bureau's records do not reveal the exact numbers in this category but only the percentage.

In processing the data, the 1970 Census used imputation techniques similar to those applied in 1960, except that the Census Bureau made far more extensive use of hot decks. It appears that the Census Bureau performed both household size and occupancy imputation in the 1970 Census. The 1970 apportionment count included about 900,000 imputed persons.

The 1980 Census made use of count imputation for household size and occupancy. As in prior censuses, the procedure was based on geographic proximity. (The specific procedures are described in Bailar (1982, pp. 7-9).) As in 1960, count imputation appears to have been used in instances of the mechanical failure of questionnaire readers. On the other hand, the sampling techniques used in 1970 in the National Vacancy Check and Post-Enumeration Post-Office Check were not continued. In 1980, approximately 761,000 persons were added to the final count through imputation (U.S. Census Bureau 2001d). After the 1980 Census, the State of Indiana sued the Census Bureau as the use of count imputation shifted a seat from Indiana to Florida (Orr v. Baldrige 1985). The suit and the precedent it set are described more in the next section.

Count imputation in the 1990 Census differed from 1980 in two major ways. First, the level of count imputation was much lower, practically nonexistent. Only about 53,600 individuals were imputed in 1990, as opposed to about 761,000 in the 1980 Census and about 900,000 individuals in 1970 (U.S. Census Bureau 1992b). The Census Bureau instituted better questionnaire control methods and more clerical edit procedures. Further, incomplete questionnaires went through additional rounds of recycling back to the field, and census workers were strongly encouraged to determine a count in the field. The result was a large drop in the number of housing units eligible for imputation. Second, and perhaps more important for this discussion, the Bureau introduced status imputation. Prior procedures had included count imputation only for household size and occupancy. The 1990 Census continued the prior practices but added status imputation.

The addition of status imputation in 1990 was required because of the introduction of a centralized, computerized address list, called the Address Control File (ACF). Address registers for prior censuses had been maintained by hand in local census offices for each separate enumeration district; no centralized address list was maintained. The 1990 ACF, in contrast, was a centralized, automated list of all housing units that was prepared in advance of the census. Its existence required reconciliation of each address having incomplete or conflicting information. If one source (the Census Control File, for example) indicated that a housing unit was a valid, nonduplicated unit, while the ACF indicated the unit should be removed from the list (deleted), this discrepancy had to be resolved. The 1990 procedures, therefore, allowed for the imputation of a status of "delete" (nonexistent), rather than only "occupied" or "vacant," to resolve discrepancies.

The Census Bureau's review of count imputation after the 1990 Census was less intense than the more comprehensive review that followed the 1980 Census. In 1990, the Bureau reasoned that further study was probably not warranted because count imputation had such a minimal effect on census accuracy (Hogan 2001). Additionally, further study of count imputation in past censuses appeared to become less relevant. As mentioned above, from the early 1990s the Bureau intended to take a sample of nonresponding housing units following the 2000 Census; such a plan would have obviated the need to conduct a separate count imputation process.

4.2 Count Imputation in the 2000 Census

With several exceptions noted below, the imputation procedures in the 2000 Census were similar to earlier ones. As in 1960, 1970, and 1980, the Census Bureau used imputation for household size and occupancy. Further, as in 1990, status imputation allowed the Bureau to impute the existence or nonexistence of a separate, non-duplicated housing unit in instances where records were not conclusive about the unit. Again, all count imputation was performed using a variant of the hot-deck method.

It is important to realize that status imputation was applied only to addresses included in the Census Bureau's Decennial Master Address File. Addresses in this file originated from official sources such as the 1990 Census, the U.S. Postal Service, local governments, tribal governments, and Census Bureau enumerators. These addresses were validated and updated according to specific procedures. No counts--based on an actual enumeration or on imputation--were attributed to housing units originating outside these procedures (Hogan 2001).

There were several criteria that led to the assignment of a delete status when the Census Bureau had information that indicated an address was a duplicate or was not valid. For example, under the Bureau's double-delete policy, a housing unit could be labeled a “vacant” or “delete” (nonexistent) only following two independent verifications. Alternatively, if the Bureau had not received a mail return from a unit and the enumerator response indicated that the unit was nonexistent, the housing unit was given a final status of delete. In all instances, housing units were only given a final status of vacant or delete whenever the questionnaire indicated in a consistent fashion that the housing unit was vacant or nonexistent (Hogan 2001).

A total of approximately 1.2 million persons, or 0.4% of the population, were added in 2000 through count imputation (U.S. Census Bureau 2001e). Although the number and percentage of count imputations were higher in 2000 than in 1990, the percentage was in line with those in earlier censuses, as shown in Table 1.

Preliminary results confirmed that most of the count imputations performed in the 2000

Census were attributable to housing units determined to exist, but whose data were not captured in the totals through a variety of processing errors (U.S. Census Bureau 2001f). These cases appear to have been appropriately included in the census. If they had not been added through count imputation, many individuals who had tried to be enumerated in the census would have been left out simply through processing errors.

Table 1. Count Imputation in the Census, 1960 - 2000

Year	Resident Population	Total Count Imputations	Percent Count Imputation
1960	179,323,175	(unknown)	(unknown)
1970	203,211,926	900,000[a]	0.44%
1980	226,545,805	761,000[a]	0.34%
1990	248,709,873	53,655	0.02%
2000	281,421,906	1,172,144	0.42%

NOTE: [a] Only approximate count imputation figures are available for 1970 and 1980.

5. UTAH CHALLENGES THE USE OF COUNT IMPUTATION

As required by the Constitution, the U.S. House of Representatives is re-apportioned after each decennial census. The Method of Equal Proportions (Ernst 1994) is used to determine how many seats are allotted to each state. In 2001 the State of Utah, finding itself fewer than 1,000 residents short of the count needed for the 435th and final seat in the House of Representatives, set about challenging the census population totals. In its first challenge, Utah alleged that its population total should have included Mormon missionaries temporarily living abroad. After this challenge was rejected by the District Court, the State brought a second complaint.

Noting that without count imputation it would be entitled to the final House seat, the State of Utah claimed that the use of count imputation violates both the Census Act and the Constitution, being merely a back-door method to implement sampling for nonresponse, which the Supreme Court had already held could not be used to produce the apportionment count. Utah sought an injunction compelling Census Bureau officials to change the official census counts. North Carolina, the state that would lose the contested seat if imputation counts were removed, intervened. After extensive briefing, the District Court found in favor of the Federal Government, Utah v. Evans (2001, Lower Court Opinion), and the State appealed directly to the Supreme Court.

5.1 Imputation Is Designed to Improve Accuracy

Utah was not the first state to challenge the Census Bureau's use of count imputation. The State of Indiana sued the Census Bureau after the 1980 Census because, without the use of count imputation, a seat would have shifted from Florida to Indiana, Orr v. Baldrige (1985). The parties stipulated that imputation was not sampling and the District Court upheld the use of imputation, holding that "the Bureau's use of hot deck imputation was an entirely reasonable means of dealing with the problem of incomplete data ..." This decision was not appealed.

In the decennial census, the basic alternatives for handling missing data for an identified address are either (1) to tabulate the address as "data not reported," that is, to assign no value to the empty data fields, or (2) to assign (impute) plausible values for the missing data. The first option is the equivalent of imputing a value of 0, that is, of deciding that all returns with questionable or incomplete data or unresolved status represent vacant or nonexistent housing units. This conclusion is demonstrably untrue, as studies have shown that a significant proportion of such returns are actually valid, occupied housing units (Hogan 2001). Under the second alternative, a value other than 0 may be inserted into the empty field. Because incomplete information may distort census results, producing an accurate census dictates that the issue is not whether to impute, but what type of imputation will be the most accurate.

In Utah v. Evans, the Census Bureau contended that imputation is recognized in the statistical community as a statistical procedure that can improve the accuracy and reliability of censuses and surveys, and that imputation has been amply discussed and documented in the professional literature for many years (see, e.g., Madow, Olkin, and Rubin 1983). The hot-deck method is recognized as a valid procedure for imputing in many circumstances, particularly in large-scale operations such as the U.S. decennial census. The assumption of local homogeneity that underlies the hot-deck technique is also well supported in practice and in the academic literature.

Imputation makes the census more accurate because refusing to impute ignores some people who tried to participate in the census. Without imputation, data from the returns of residents of housing units with conflicting or missing data would be excluded from the census, despite their efforts to be included. Even extensive attempts to obtain direct information from households will occasionally fail; in those instances, imputation permits their occupants to be accounted for in the census.

5.2 Imputation Is Distinguishable From Sampling

Expert witnesses testifying for the State of Utah contended that imputation was a form of sampling, which was prohibited by the Supreme Court in Dept. of Commerce v. U.S. House of Representatives (1999). Quoting from Last (1995, p. 151), Lara Wolfson defined sampling as "the process of selecting a number of subjects [units] from all the subjects [units] in a particular group or 'universe'." She added that "[i]n contemporaneous usage ... 'statistical sampling' is taken not only to be the process of selecting the subjects, but the subsequent use of the sample to draw inference to the population from which it was drawn" (Wolfson 2001, p. 13).

This definition does not apply to the 2000 Census, where the Census Bureau attempted in numerous ways to contact every resident and made no "selection" of a subgroup in designing the census. Further, it does not make clear that, in sampling, the process of selecting a sample is a deliberate and purposeful activity occurring during the design of a survey. Without this understanding, the definition is broad enough to cover situations that have nothing to do with sampling; with this understanding, the definition does not encompass imputation. Imputation is not a mechanism for selecting units during the data collection stage of a census or sample survey, but rather a means of addressing missing data in the data processing stage.

Donald Rubin, also writing for the State of Utah, defined sampling as follows: "Sampling refers to the process of obtaining data from a subset of a population (the subset is usually called the 'sample') from which estimates are made about characteristics of the entire population" (Rubin, p. 4). The argument is that the pre-imputation values, that is, the population counts from enumerated households, constitute a sample. This "sample" is then used, through the imputation process, to estimate the values (population counts) for the whole population.

The second definition suffers from the same flaws as the first in that it does not incorporate the process of *deliberately selecting* a subset of the population *during the design phase of a survey*. Indeed, Dr. Rubin apparently recognized this problem with his definition, because later in his declaration he writes that the "way in which the sample is chosen is critical" and that "[i]n order to use a sample for valid scientific estimation, specific sampling design requirements must be adhered to" (p. 4). Without the notion of deliberate or purposeful selection according to a sampling design, however, the second definition is as overly broad as the first. *Under these definitions, all censuses in this or probably any other country have been "samples," as none to our knowledge have ever included each and every person in the population.*

Dr. Wolfson also claimed that imputation and the methodology struck down in Dept. of Commerce v. U.S. House of Representatives are both statistical sampling because both "attempt to estimate the number of persons in households that were not actually enumerated by traditional census methods on the basis of statistical methods and assumptions that allow inferences to be drawn about an unobserved segment of the population based on information in an observed segment of the population" (p. 14). Essentially, this says that sampling is a process that allows inferences to be drawn about an unobserved segment of the population based on information in an observed segment of the population. As will be explained in the subsequent paragraphs, this definition is entirely too broad.

There are various kinds of sampling, including probability and nonprobability sampling. The lead witness for the Census Bureau, Howard Hogan, contended that when used in the context of statistical, demographic, and population survey applications, the term "sampling" generally refers to probability sampling. Dr. Hogan cited testimony from Barbara Bailar, formerly the Associate Director for Statistical Standards and Methodology at the Census Bureau, in Orr v. Baldrige in arguing that sampling is "the selection of a subset of units from a larger population in such a way that each unit of the population has a known chance of selection. Sampling is used where a scientifically selected set of units can be used to represent the entire population from which they are drawn and inferences to the entire population can be based on sample results" (Bailar 1982, p. 5).

There are other forms of sampling. Sampling in connection with the census (that is, the asking of certain questions to only a fraction of the population) began in 1940 with probability sampling. As Dr. Rubin conceded, the imputation used in the 2000 Census was not probability sampling because the methodology was unrelated to the selection of a subset of units from a larger population in such a way that each unit of the population has a known chance of selection.

The Census Bureau's argument follows. In the Census Act, Congress ordered the Bureau not to use "the statistical method known as 'sampling,'" with "sampling" in quotation marks. Congress was referring to what statisticians normally meant by sampling at the time the Census Act was amended in 1957 to include the word sampling. Sukhatme (1954, p. 1) provided an accepted definition: "A sampling method is a method of selecting a fraction of the population in a way that the selected sample represents the population." The Sukhatme definition represents what statisticians meant by sampling in 1957 and today. This definition is consistent with the definition used by the Supreme Court in both Utah v. Evans and the earlier Dept. of Commerce v. U.S. House of Representatives. As normally used in the statistical profession, "the statistical method know as 'sampling'" is the deliberate selection of some units, usually a small fraction.

In its declarations, the Census Bureau tried to make clear that (1) various methodologies are statistical procedures, including sampling, editing, imputation, weighting, quality control, etc.; and (2) imputation is one of many statistical methodologies, but it is not sampling.

A review of how sampling and imputation relate, through an examination of the entirety of data collection and processing, demonstrates that sampling and imputation are not the same. When one wants to produce data describing a target population or subgroup, a set of activities is conducted that can be split into two stages: (1) collecting the data, and (2) processing the data. Both collection and processing can similarly be broken into smaller parts. Sampling may or may not be used in the first stage (data collection), while imputation is only one of several activities in the second stage (processing). The uses and purposes of imputation are therefore distinct from those of sampling.

To see this more clearly, when the Census Bureau conducts a census or a survey, it usually proceeds through the following two-stage operation.

> *Stage 1. Data Collection:* First, the Census Bureau decides which units it will attempt to contact, that is, every unit in the target group, or only a sample. If the Bureau is taking a sample survey, it then determines the sample design, that is, how to select the people or units to be contacted. Other aspects determined at this stage include the following: the time period for collecting the data, the questionnaire to be used, how the survey or census takers will deal (in the field) with nonrespondents, etc. Generally, the Bureau prepares a list of the units eligible to be contacted (the frame) and, then, if a sample is being taken, the units to be included in the sample survey are selected from this frame according to the sample design. In a census, all units in the frame are included. The data are then collected through interviews, either in person, over the telephone, via mail, or through other modes. Finally, for most censuses and sample surveys, follow-up efforts must be made to recontact people who do not respond initially. These efforts might entail visiting a person who did not reply by mail or making follow-up telephone calls.
>
> *Stage 2. Data Processing:* The second stage is data processing. Whether one is dealing with a census or a survey, the data already collected are put into a centralized database. Then, the data are processed and edited through a variety of methods. One such method, imputation, is used to assign a value to an empty field on a partially completed questionnaire, or to complete a blank questionnaire. The purpose is to form a complete record for a given respondent. The imputation method selected may rely on prior experience or on data already collected to replace the missing data. Editing also occurs at this stage for responses that fail checks to ensure that the data on the entire questionnaire are consistent. (For example, the response could indicate that a person is two years old and also the spouse of the householder.) The inconsistent or incomplete item is replaced

> with an imputed value that is consistent with other responses on the questionnaire. The data are then tabulated to determine totals, means, proportions, ratios, or other statistics. For sample surveys, weights that incorporate the probability of selection and possibly other adjustments are usually used in these computations.

Thus, the term sampling refers to one strategy of data collection and can be said to encompass some or all of the activities in the data collection stage of, for example, a sample survey. By contrast, imputation--count imputation or other types--is a statistical procedure applied only in the second stage, the data processing stage. Bailar stated in her affidavit that sampling and imputation "are two completely different procedures, based upon totally distinct principles and serving equally distinct purposes" (Bailar 1982, p. 5). Expanding, she said that sampling and imputation are not competitive, nor is one methodology superior to the other. As Bailar demonstrates, the Bureau's longstanding position has been that imputation and sampling are different.

Joseph Waksberg, an expert witness writing for the Census Bureau in Utah v. Evans, also made the point that (1) research topics in surveys and censuses are very often classified into two categories--sampling and nonsampling; and (2) "[m]ost survey researchers and practitioners consider nonresponse and methods of adjusting for nonresponse either by imputation, as it has been done in recent decennial censuses, or through weighting, which is applied more commonly in sample surveys, as nonsampling issues and treat them as separate and distinct from sampling issues" (Waksberg 2001, p. 4). To demonstrate his point, Mr. Waksberg went on to provide a variety of supporting examples, including a short pamphlet, *What is a Survey?*, printed by the American Statistical Association (ASA 1980), a three-volume report published by the Committee on National Statistics (Madow et al. 1983), a book on nonsampling errors in surveys (Lessler and Kalsbeek 1992), and a conference on telephone survey methodology.

In Utah v. Evans, the Supreme Court in a split decision agreed with the Census Bureau that imputation was not sampling. The majority held that "imputation differs from sampling in respect to the nature of the enterprise, the methodology used, and the immediate objective sought." The Court further noted that the differences were "of both kind and degree."

Justice Breyer, writing for the majority, summarized well the distinction between sampling and imputation, citing with approval an illustrative example offered in oral argument:

> Imagine a librarian who wishes to determine the total number of books in a library. If the librarian finds a statistically sound way to select a sample (*e.g.*, the books contained on every 10th shelf) and if the librarian then uses a statistically sound method of extrapolating from the part to the whole (*e.g.*, multiplying by 10), then the librarian has determined the total number of books by using the statistical method known as "sampling." If, however, the librarian simply tries to count every book one by one, the librarian has not used sampling. Nor does the latter process suddenly become "sampling" simply because the librarian, finding empty shelf spaces, "imputes" to that empty shelf space the number of books (currently in use) that likely filled them--not even if the librarian goes about the imputation process in a rather technical way, say by measuring the size of nearby books and dividing the length of each empty shelf space by a number

> representing the average size of nearby books on the same shelf. ... *[citations omitted]* (Utah v. Evans 2002, Opinion of the Court, p. 10)

The Court recognized that sampling and imputation differ in three critical respects:

> (1) In respect to the *nature of the enterprise,* the librarian's sampling represents an overall approach to the counting problem that from the beginning relies on data that will be collected from only a part of the total population; ... (2) in respect to *methodology,* the librarian's sampling focuses on using statistically valid sample-selection techniques to determine what data to collect; ... and (3) in respect to *the immediate objective,* the librarian's sampling seeks immediately to extrapolate the sample's relevant population characteristics to the whole population. ...
>
> By way of contrast, the librarian's imputation (1) does not represent an overall approach to the counting problem that will rely on data collected from only a subset of the total population, since it is a method of *processing* data (giving a value to missing data), not its collection; ... it (2) does not rely upon the same statistical methodology generally used for sample selection; ... and it (3) has as its immediate objective determining the characteristics of missing individual books, not extrapolating characteristics from the sample to the entire book population. ... *[citations omitted]* (pp. 10-11)

The Court agreed that count imputation was not used for "the extrapolation of the features of a large population from a small one, but the filling in of missing data as part of an effort to count individuals one by one" (p. 11). The Court further examined the legislative history and wording of the Census Act, as well as the definition of "sampling" at the time the Act was amended, to conclude that Congress did not intend to bar the use of all statistical methods, just sampling.

Justice O'Connor dissented from the holding, concluding that the imputation did constitute a form of sampling and was thus barred by the Census Act. Justice Scalia dissented on the grounds that the Court did not have jurisdiction to hear the case. Justices Thomas and Kennedy agreed that imputation was distinguishable from sampling, but would have found its use unconstitutional. These two justices concluded that the Framers of the Constitution had experience with various statistical techniques and purposefully disallowed them in the census.

5.3 Imputation Does Not Violate the Constitutional Requirement of an "actual Enumeration"

The majority of the Court, however, held that the use of count imputation did not violate the Constitution's requirement of an "actual Enumeration." Utah had argued that the actual Enumeration requirement compelled the Census Bureau to seek out each individual. The Census Bureau could rely on documentary or eyewitness evidence that an individual existed, and it could rely on proxy respondents. But, Utah contended, the Census Bureau could not rely on count imputation to fill in missing data. The Supreme Court examined the intent of the Framers and contemporaneous usage of the word enumeration before concluding that count imputation, as used in the 2000 Census, did not violate the Constitution.

The Supreme Court upheld the use of count imputation, noting that the interest in accuracy favored the Census Bureau. Also important in the decision was the fact that count imputation was used only as a last resort, after all other methods had failed. The Court did not uphold the use of all forms and uses of count imputation or other statistical methodologies. Rather the Court approved the use of count imputation in the 2000 Census after all efforts were made to reach each household, where a very small portion (less than 0.5%) of the population was accounted for with count imputation, where the alternative to count imputation was decreased accuracy, and where manipulation of the method was highly unlikely.

6. THE FUTURE

The Supreme Court recognized in Utah v. Evans that the Constitution is a flexible document, a document capable of confronting future changes in society and technology, noting that,

> ... however unaware the Framers might have been of specific future census needs, say, of automobiles for transport or of computers for calculation, they fully understood that those future needs might differ dramatically from those of their own times. And they were optimists who might not have been surprised to learn that a year 2000 census of the Nation that they founded required 'processed data for over 120 million households, including over 147 million paper questionnaires and 1.5 billion pages of printed material.'

Utah v. Evans is a recognition that the Constitution can accommodate changes in census methodology that are consistent with the goal of equal representation. The false dichotomy of statistics and an "actual Enumeration" will not stand; rather future statistical improvements to census methodology must be judged by the methodology used, the kind and degree of statistical inference, and the immediate objective of the procedure. The boundaries of Utah v. Evans are not exact. This lack of definitiveness is appropriate given the changing nature of census methods and statistical advances. The Census Bureau will continue to propose and test future census taking methods, statistical and nonstatistical, to plan for the most accurate census possible within the bounds of resources and respondent burden, and consistent with the law and the Constitution. There is an appropriate role for Congress in the decision-making processes for future censuses. The use of statistics must evolve in a manner consistent with Congressional understanding and direction.

A census is necessarily an interaction between technology and society. American society is constantly changing. Language use, ethnic composition, household structure, and housing patterns constantly change. New communities are built seemingly overnight; old communities evolve or slowly fade. Attitudes towards privacy and the appropriate role of government shift. Factors such as cell phones, gated communities, and long commutes affect household accessibility.

By many measures, the 2000 Census adapted successfully to this changing landscape. The estimated net coverage error was quite small. The revised demographic analysis measured a 0.1% net undercount; the Accuracy and Coverage Evaluation Revision II estimated a net *overcount* of about 0.5%. Additionally, the 2000 Census is believed to have reduced, but not eliminated, the differential net undercount, the historical phenomenon of undercounting minority populations and untraditional households (U.S. Census Bureau 2003, p. ii). These commendable results were achieved in spite of the changing landscape and with the use of statistical methods.

Turning to the future, the Census Bureau Director, C. Louis Kincannon, has stated:

> We are not planning, however, to request additional funding for 2010 to further develop this methodology to adjust the census for the purpose of redistricting. My opinion, from talking with the Census Bureau staff, is that we do not have a methodology that will produce accurate numbers for redistricting in the proper time frame. We know that the likelihood of success there, does not justify the expense when there are other demands that *must* be met, including coverage improvement; updating the Master Address File and the maps; modernizing the data collection operations; expanded language programs; and other operations that will improve the Census Bureau's ability to count the people of your cities.
>
> We will always bear in mind that the decennial census belongs to the people of the United States and we will seek the counsel of our stakeholders, especially our peers in the scientific and policy community, as we plan for the future. (C. Louis Kincannon, Director, U.S. Census Bureau, remarks before the U.S. Conference of Mayors, January 22, 2003)

Adjustment is but one of many statistical tools. The census in 2010 and beyond will confront new and unforeseen challenges as well as many familiar ones. To meet those challenges, census planners will have at their disposal a wide variety of tools, including the tools of modern statistical science.

REFERENCES

Abstract of the Seventh Census (1852), *The Seventh Census. Report of the Superintendent of the Census for December 1, 1852,* Washington, D.C.: Printed by Order of the House of Representatives of the U.S.

American Statistical Association (1980), *What is a Survey?*, Alexandria, Va.: Author.

Anderson, M. (1988), *The American Census: A Social History*, New Haven, Conn.: Yale University Press.

Anderson, M. J. and Fienberg, S. E. (1999), *Who Counts? The Politics of Census-Taking in Contemporary America*, New York, N.Y.: Russell Sage Foundation.

Bailar, B. (1982), "Affidavit of Barbara A. Bailar," Orr v. Baldrige, U.S. District Court for the Southern District of Indiana, Indiana Division, Civil Action No. IP-81-604C, February 19, 1982.

Coale, A. J. (1955), "The Population of the United States in 1950 Classified by Age, Sex, and Color--A Revision of Census Figures," *Journal of the American Statistical Association*, *50*, 16-54.

Department of Commerce, "Adjustment of the 1990 Census for Overcounts and Undercounts of Population and Housing: Notice of Final Decision," Federal Register (July 22, 1991), vol. 56, 33583.

Ernst, L. R. (1994), "Apportionment Methods for the House of Representatives and the Court Challenges," *Management Science, 40*, 1207-1227.

Hogan, H. (2001), "Declaration of Howard Hogan," Utah v. Evans, U.S. District Court for the District of Utah, Central Division, Civil No. 2001:CV-000292G, July 11, 2001. An abridged version is available at *Proceedings of the Annual Meeting of the American Statistical Association*, New York, N.Y., 2002, Section on Survey Methods Research, 1475-1484.

Hogan, H. and Waite, P. J. (1998), "Statistical Methodologies for Census 2000: Decisions, Issues, and Preliminary Results," *Proceedings of the Social Statistics Section, American Statistical Association*, 40-49.

Hogan, H. and Wolter, K. (1988), "Measuring Accuracy in a Post-Enumeration Survey," *Survey Methodology*, 14, 99-116.

Last, J. M. (ed.) (1995), *A Dictionary of Epidemiology, 3rd Ed.*, Oxford, United Kingdom: Oxford University Press.

Lessler, J. T. and Kalsbeek, W. D. (1992), *Nonsampling Errors in Surveys*, New York, N.Y.: J. Wiley and Sons, Inc.

Madow, W. G., Olkin, I., and Rubin, D. (eds.) (1983), *Incomplete Data in Sample Surveys, Volume 2, Theory and Bibliographies*, New York, N.Y.: Academic Press.

National Research Council (1995), *Modernizing the U.S. Census,* Panel on Census Requirements in the Year 2000 and Beyond, Washington, D.C.: National Academy Press.

Price, D. O. (1947), "A Check on Underenumeration in the 1940 Census," *American Sociological Review, Vol. 12*, Issue 1 (Feb.), 44-49.

Rubin, D. (2001), “Declaration of Donald Rubin,” Utah v. Evans, U.S. District Court for the District of Utah, Central Division, Civil No. 2001:CV-000292G, April 12, 2001.

Sekar, C. C. and Deming, W. E. (1949), “On a Method of Estimating Birth and Death Rates and the Extent of Registration,” *Journal of the American Statistical Association*, *44*, 101-115.

Sukhatme, P. V. (1954), *Sampling Theory of Surveys With Applications*, Bangalore, India: The Bangalore Press.

U.S. Census Bureau (1965), *Censuses of Population and Housing: Procedural History*, Washington, D.C.: U.S. Department of Commerce.

---------- (1967), “Quality Control of the Field Enumeration,” 1960 Census of Population and Housing, Washington, D.C.

---------- (1971a), “Results of National Vacancy Check,” 1970 Census Preliminary Evaluation Results Memorandum No. 12, Memorandum from B. Denton for M. Boisen, July 27, 1971.

---------- (1971b), “Final Tally of Census Adds in Sixteen Southern States,” 1970 Census Preliminary Evaluation Results Memorandum No. 11, Memorandum from B. Denton for M. Boisen, July 27, 1971.

---------- (1975), *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition, Part 2*, Washington, D.C.: U.S. Government Printing Office.

---------- (1984), "Preliminary Summary of Results from the 1980 Census Coverage Improvement Program Evaluations," Memorandum from J. H. Thompson for C. D. Jones, July 19, 1984.

---------- (1992a), "Final Results from the Parolee/Probationer Coverage Improvement Program and the Parolee/Probationer Coverage Improvement Followup Program," Memorandum from J. H. Thompson for C. D. Jones, August 10, 1992.

---------- (1992b), “Imputation in the 1990 Census - Coverage Implications,” DSSD REX 1990 Memorandum Series #BB-7, Memorandum from D. Griffin for C. D. Jones, September 4, 1992.

---------- (1997), “Report to Congress--The Plan for Census 2000,” Washington, D.C.: U.S. Department of Commerce.

---------- (2000a), "Census 2000 Service Based Enumeration: Overview of Multiplcity Estimation," DSSD Census 2000 Procedures and Operations Memorandum Series #Q-36, Memorandum from D. Kostanich for Howard Hogan, December 5, 2000.

---------- (2000b), “Overview of the Duplicate Housing Unit Operations,” Census 2000

Informational Memorandum No. 78, Memorandum from S. Miskura to the Distribution List, November 7, 2000.

---------- (2000c), "Results of Reinstatement Rules for the Housing Unit Duplication Operations," Census 2000 Informational Memorandum No. 82, Memorandum from S. Miskura to the Distribution List, November 21, 2000.

---------- (2001a), "Census 2000 - Service Based Enumeration Multiplicity Estimation," DSSD Census 2000 Procedures and Operations Memorandum Series B-15*, Memorandum from D. Kostanich for H. Hogan, February 28, 2001. On-line: http://www.census.gov/dmd/www/pdf/Fr15.pdf.

---------- (2001b), "Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy: Recommendation Concerning the Methodology to be Used in Producing the Tabulations of Population Reported to States and Localities Pursuant to 13 U.S.C. 141(c)," March 1, 2001, Washington, D.C. On-line: http://www.census.gov/dmd/www/pdf/Escap2.pdf.

---------- (2001c), "Report of the Executive Steering Committee for Accuracy and Coverage Evaluation Policy on Adjustment for Non-Redistricting Uses," October 17, 2001, Washington, D.C. On-line: http://www.census.gov/dmd/www/pdf/Recommend2.pdf.

---------- (2001d), "Comparison of Record Substitution and 100% Data Allocation: 1980, 1990, and 2000 Censuses," Memorandum from S. Love and D. Dalzell for the Record, March 27, 2001.

---------- (2001e), "Census 2000 - Missing Housing Unit Status and Population Data," DSSD 2000 Procedures and Operations Memorandum Series B-17*, Memorandum from D. Kostanich for H. Hogan, February 28, 2001. On-line: http://www.census.gov/dmd/www/pdf/Fr17.pdf.

---------- (2001f), "Initial Research on Count Imputation in Census 2000," Census 2000 Informational Memorandum No. 110, Memorandum from T. Angueira for P. J. Waite, August 10, 2001.

---------- (2003), "Technical Assessment of A.C.E. Revision II," March 12, 2003. On-line: http://www.census.gov/dmd/www/pdf/ACETechAssess.pdf.

Waksberg, J. (2001), "Declaration of Joseph Waksberg," Utah v. Evans, U.S. District Court for the District of Utah, Central Division, Civil No. 2001:CV-000292G, July 10, 2001. An abridged version is available at *Proceedings of the Annual Meeting of the American Statistical Association*, New York, N.Y., 2002, Section on Survey Methods Research, 3577-3578.

Wolfson, L. (2001), "Declaration of Lara J. Wolfson, Ph.D." Utah v. Evans, U.S. District Court for the District of Utah, Central Division, Civil No. 2001:CV-000292G, June 10, 2001. An abridged version is available at *Proceedings of the Annual Meeting of the American Statistical*

Association, New York, N.Y., 2002, Section on Survey Methods Research, 3762-3771.

COURT CASES CITED

Heard in the U.S. Supreme Court:

Utah v. Evans, 526 U.S. 425 (2002).

Department of Commerce v. United States House of Representatives, 525 U.S. 316 (1999).

Wisconsin v. City of New York, 517 U.S. 1 (1996).

Franklin v. Massachusetts, 505 U.S. 788 (1992).

Heard in U.S. District Court:

Utah v. Evans, 182 F. Supp. 2d 1165, D. Utah (2001).

Orr v. Baldrige (unreported), No. IP-81-604-C, slip op., D. Ind. (1985).